



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **10049414 A**(43) Date of publication of application: **20.02.98**

(51) Int. Cl. **G06F 12/00**
G06F 12/00
G06F 12/00
G06F 17/30

(21) Application number: **08206145**(22) Date of filing: **05.08.96**(71) Applicant: **TOSHIBA CORP**

(72) Inventor: **SAWAJIMA SHINSUKE**
SUZUOKA SETSU
YAMANE TETSUYA

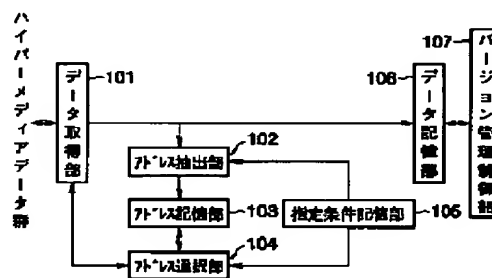
(54) **DEVICE AND METHOD FOR VERSION
MANAGEMENT**

(57) Abstract:

PROBLEM TO BE SOLVED: To manage the versions of respective data only by irreducible specification by adding its address and version information regarding its version to acquired data, etc., and storing them.

SOLUTION: A data acquisition part 101 acquires data of a specific address from a hypermedia data group. An address extraction part 102 extracts addresses embedded as anchors in the data and stores addresses which are stored in a specific condition storage part 105 and meet specific conditions in an address storage part 103. An address selection part 104 indicates addresses to be acquired from the stored addresses. A specific condition storage part 105 stores specific conditions such as address acquisition time intervals of addresses, etc. A data storage part 106 adds addresses and version information to the acquired data or their processed data and stores them. A version management control part 107 performs control regarding the version management by using the stored information.

COPYRIGHT: (C)1998,JPO



(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平10-49414

(43)公開日 平成10年(1998)2月20日

(51)Int.Cl. ⁸	識別記号	庁内整理番号	FI	技術表示箇所
G 0 6 F 12/00	5 1 7		G 0 6 F 12/00	5 1 7
	5 4 5			5 4 5 Z
	5 4 7			5 4 7 H
17/30			15/401	3 4 0 C

審査請求 未請求 請求項の数7 O L (全 13 頁)

(21)出願番号 特願平8-206145

(22)出願日 平成8年(1996)8月5日

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 澤島 信介

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

(72)発明者 鈴木 節

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

(72)発明者 山根 徹也

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

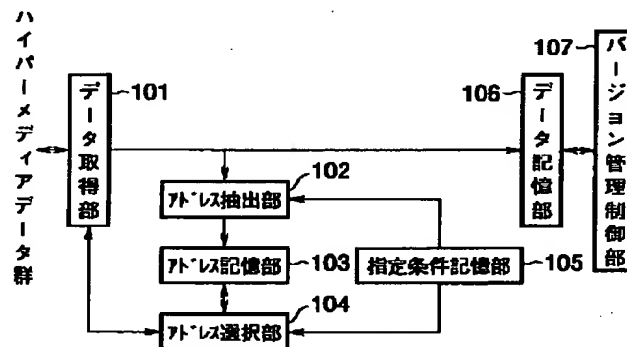
(74)代理人 弁理士 鈴木 武彦 (外6名)

(54)【発明の名称】 バージョン管理装置及びバージョン管理方法

(57)【要約】

【課題】 個々のハイパーメディアデータすべてを指定せず最低限の指定のみで各データを取得しバージョンを管理するバージョン管理装置を提供すること。

【解決手段】 アドレス指定により取得するハイパーメディアデータのバージョンを管理する装置であつて、指定のアドレスのハイパーメディアデータの取得を試行するデータ取得手段と、取得されたハイパーメディアデータに埋め込まれているアドレスを抽出する手段と、抽出されたアドレスを記憶する手段と、この手段からアドレスを選択し、該アドレスで指定されるデータの取得をデータ取得手段に指示する手段と、データ取得手段により取得されたデータまたは該データを加工したものに該データのアドレスと該データのバージョンに関する情報とを付加して記憶する手段とを備えたことを特徴とする。



【特許請求の範囲】

【請求項 1】アドレスの指定により取得可能なハイパーメディアデータのバージョンを管理するバージョン管理装置であって、

指定されたアドレスのハイパーメディアデータの取得を試行するデータ取得手段と、

このデータ取得手段により取得されたハイパーメディアデータにアンカーとして埋め込まれているアドレスを抽出するアドレス抽出手段と、

少なくとも前記アドレス抽出手段により抽出されたアドレスを記憶するアドレス記憶手段と、

このアドレス記憶手段からアドレスを選択し、該アドレスで指定されるデータの取得を前記データ取得手段に指示するアドレス選択手段と、

前記データ取得手段により取得されたデータまたは該データを加工したものに、該データのアドレスと、該データのバージョンに関するバージョン情報とを付加して記憶するデータ記憶手段とを備えたことを特徴とするバージョン管理装置。

【請求項 2】前記アドレス記憶手段には、前記アドレスに該アドレスで指定されるデータの取得の試行に関するアドレス付加情報を付加して記憶し、

前記アドレス選択手段は、前記アドレス付加情報を参照して前記アドレス記憶手段からアドレスを選択することを特徴とする請求項 1 に記載のバージョン管理装置。

【請求項 3】ユーザからの要求に応じて前記データ記憶手段に記憶されているデータを検索する検索サービス手段をさらに備えたことを特徴とする請求項 1 に記載のバージョン管理装置。

【請求項 4】ユーザからの要求に応じて前記データ記憶手段から指定されたバージョンのデータを取り出すバージョン管理制御手段をさらに備えたことを特徴とする請求項 1 に記載のバージョン管理装置。

【請求項 5】アドレス単位で指定可能な、データ取得試行対象およびデータ取得試行タイミングの少なくとも一方に関する指定条件を記憶する指定条件記憶手段をさらに備え、

前記アドレス選択手段は、前記指定条件記憶手段に記憶された前記指定条件に基づいて、前記アドレス記憶手段からアドレスを選択することを特徴とする請求項 1 に記載のバージョン管理装置。

【請求項 6】前記アドレス抽出手段により抽出されたアドレスのうち、前記指定条件記憶手段に記憶された前記指定条件に基づいて選択したアドレスのみを前記アドレス記憶手段に記憶させることを特徴とする請求項 5 に記載のバージョン管理装置。

【請求項 7】アドレスの指定により取得可能なハイパーメディアデータのバージョンを管理するバージョン管理方法であって、

選択されたアドレスのハイパーメディアデータの取得を

試行し、

取得されたハイパーメディアデータにアンカーとして埋め込まれているアドレスを抽出し、抽出されたアドレスを記憶するとともに、

取得されたデータまたは該データを加工したものに、該データのアドレスと、該データのバージョンに関するバージョン情報とを付加して記憶し、

記憶された前記アドレスのうち所定の選択条件を満たすものを選択して、該アドレスで指定されるデータの取得を指示することを特徴とするバージョン管理方法。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】本発明は、アドレスの指定により取得可能なハイパーメディアデータのバージョンを管理するバージョン管理装置及びバージョン管理方法に関する。

【0002】

【従来の技術】一般にデータは作成された後、更新されたり削除されたりするため、過去のバージョンを記憶させ、後で取り出したり、新バージョンと比較したりするバージョン管理が行われてきた。そして、従来のバージョン管理ツールでは、管理の対象となる個々のデータをあらかじめ指定しておく必要があった。

【0003】ところで、データの個数が多かったり、それらがネットワークを介して分散していたり、さらには複数の人によってデータの追加、削除、更新が行われたりすると、個々のデータの状態を把握することが難しくなる。従って、結果的に管理したい個々のデータを指定することが難しくなり、バージョン管理ができなくなるという問題点があった。特に、ハイパーメディアデータのバージョン管理においては、各ハイパーメディアデータについて単に一般的なデータと同様の管理の仕方を行うバージョン管理ツールがあるだけで、ハイパーメディアであることの特性を利用してバージョンを管理するのは存在しなかった。

【0004】一方、多数のデータの中から必要なデータを捜し出すには、全データを対象として検索を行うことが有効である。ハイパーメディアの大規模なものについては、例えばインターネットにおけるWWW (World Wide Web) 上のハイパーメディアデータを対象とした検索サービスがすでに存在する。しかしながら、このような検索サービスはデータの比較的新しい唯一のバージョンのみを対象としており、過去のバージョンをも対象とする検索サービスは存在しなかった。

【0005】

【発明が解決しようとする課題】従来はハイパーメディアデータのバージョン管理を行う場合でも、個々のデータを指定する必要があり、データ数が多いときなどはバージョン管理が困難になるという問題があった。

【0006】また、ハイパーメディアデータを対象とし

た検索において、過去のバージョンを対象とした検索ができないという問題もあった。本発明は、上記事情を考慮してなされたものであり、バージョン管理の対象とすべき個々のハイパーメディアデータすべてを指定することを必要とせず、最低限の指定のみで、各データを取得し夫々のデータのバージョンを管理するバージョン管理装置及びバージョン管理方法を提供することを目的とする。

【0007】

【課題を解決するための手段】本発明（請求項1）は、アドレスの指定により取得可能なハイパーメディアデータのバージョンを管理するバージョン管理装置であって、指定されたアドレスのハイパーメディアデータの取得を試行するデータ取得手段と、このデータ取得手段により取得されたハイパーメディアデータにアンカーとして埋め込まれているアドレスを抽出するアドレス抽出手段と、少なくとも前記アドレス抽出手段により抽出されたアドレス夫々を記憶するアドレス記憶手段と、このアドレス記憶手段からアドレスを選択し、該アドレスで指定されるデータの取得を前記データ取得手段に指示するアドレス選択手段と、前記データ取得手段により取得されたデータまたは該データを加工したものに、該データのアドレスと、該データのバージョンに関するバージョン情報とを付加して記憶するデータ記憶手段とを備えたことを特徴とする。

【0008】本発明（請求項2）は、請求項1に記載のバージョン管理装置において、前記アドレス記憶手段には、前記アドレスに該アドレスで指定されるデータの取得の試行に関するアドレス付加情報を付加して記憶し、前記アドレス選択手段は、前記アドレス付加情報を参照して前記アドレス記憶手段からアドレスを選択することを特徴とする。

【0009】本発明（請求項3）は、請求項1に記載のバージョン管理装置において、ユーザからの要求に応じて前記データ記憶手段に記憶されているデータを検索する検索サービス手段をさらに備えたことを特徴とする。

【0010】本発明（請求項4）は、請求項1に記載のバージョン管理装置において、ユーザからの要求に応じて前記データ記憶手段から指定されたバージョンのデータを取り出すバージョン管理制御手段をさらに備えたことを特徴とする。

【0011】また、本発明は、請求項1に記載のバージョン管理装置において、前記バージョン情報は、データの取得時刻、データの記憶時刻、データのサイズおよび同一アドレスのものについて何回目に記憶されたかを示すバージョン番号の少なくとも1つを含むようにしても良い。

【0012】本発明（請求項5）は、請求項1に記載のバージョン管理装置において、アドレス単位で指定可能な、データ取得試行対象およびデータ取得試行タイミン

グの少なくとも一方に関する指定条件を記憶する指定条件記憶手段をさらに備え、前記アドレス選択手段は、前記指定条件記憶手段に記憶された前記指定条件に基づいて、前記アドレス記憶手段からアドレスを選択することを特徴とする。

【0013】前記指定条件は、例えば、データの取得を開始するアドレスである起点アドレス、アドレスで指定されるデータの取得予定時間または取得時間間隔、データを取得するアドレスの範囲などである。

10 【0014】アドレスで指定されるデータの取得予定時間をもとにしてアドレスを選択する場合、例えば、現在時刻を参照し、取得予定時間となったアドレスを選択する。また、アドレスで指定されるデータの取得時間間隔をもとにする場合、例えば、該アドレスに付加された前記アドレス付加情報から前回に取得した時刻を調べ、これと現在時刻とを比べることにより、該当するに至ったアドレスを選択する。また、例えば、前記アドレス記憶手段に記憶されているアドレスであっても、アドレスデータを取得するアドレスの範囲に該当しないものは選択しない。

20 【0015】本発明（請求項6）は、請求項5に記載のバージョン管理装置において、前記アドレス抽出手段により抽出されたアドレスのうち、前記指定条件記憶手段に記憶された前記指定条件に基づいて選択したアドレスのみを前記アドレス記憶手段に記憶させることを特徴とする。

30 【0016】また、本発明は、前記データ記憶手段は、すでに記憶されているデータと同一アドレスのデータを記憶するときに、旧データを他の記憶されているデータにより再現できる形に加工するかまたはそのままの状態

40 【0017】また、本発明は、前記データ記憶手段は、すでに記憶されているデータと同一アドレスのデータを記憶するときに、新データを記憶する代わりに、旧データ（直前のバージョン）と新データとの差分を記憶するようにしても良い。

【0018】また、本発明は、前記データ記憶手段は、すでに記憶されているデータと同一アドレスのデータを記憶するときに、旧データ（直前のバージョン）を新データとの差分に置き換えて記憶し直し、新データを追加して記憶するようにしても良い。

50 【0019】また、本発明は、前記検索サービス手段は、ユーザから指定された時刻範囲のデータを対象として検索を行うようにしても良い。本発明（請求項7）は、アドレスの指定により取得可能なハイパーメディアデータのバージョンを管理するバージョン管理方法であって、選択されたアドレスのハイパーメディアデータの取得を試行し、取得されたハイパーメディアデータにアンカーとして埋め込まれているアドレスを抽出し、抽出

されたアドレスを記憶するとともに、取得されたデータまたは該データを加工したものに、該データのアドレスと、該データのバージョンに関するバージョン情報とを付加して記憶し、記憶された前記アドレスのうち所定の選択条件を満たすものを選択して、該アドレスで指定されるデータの取得を指示することを特徴とする。

【0020】本発明によれば、取得されたハイパーメディアデータにアンカーとして埋め込まれているアドレスにより指示されるハイパーメディアデータを抽出して連鎖的に取得対象としていくので、バージョン管理すべきハイパーメディアデータのアドレスすべてを指定することなく、一部のアドレスを指定するだけで、次々にハイパーメディアデータを取得することができるとともに、所定のタイミングで繰り返し各ハイパーメディアデータを取得することができ、また、各タイミングで取得された各ハイパーメディアデータについてのバージョン管理が可能となる。

【0021】これにより、各ハイパーメディアデータについて様々なバージョンの検索が可能となる。なお、以上の各装置に係る発明は、方法に係る説明としても成立する。

【0022】ところで、上記の発明は、相当する処理を実行させるためのプログラムを格納した記憶媒体としても成立する。例えば、本発明は、選択されたアドレスのハイパーメディアデータの取得を試行させ、取得されたハイパーメディアデータにアンカーとして埋め込まれているアドレスを抽出させ、抽出されたアドレスを記憶させるとともに、取得されたデータまたは該データを加工したものに、該データのアドレスと、該データのバージョンに関するバージョン情報とを付加して記憶させ、記憶された前記アドレスのうち所定の選択条件を満たすものを選択させて、該アドレスで指定されるデータの取得を指示させるようにコンピュータを制御するためのプログラムを格納したコンピュータ読取可能な記録媒体である。

【0023】

【発明の実施の形態】以下、図面を参照しながら発明の実施の形態を説明する。まず、本実施形態で用いる語句について説明する。「ハイパーメディアデータ」とは、データ自身の中に他のデータまたは自分自身のデータを指すアドレスが0または1以上埋め込まれているデータである。また、このように埋め込まれたアドレスを「アンカー」と呼ぶ。

【0024】なお、アンカーが0であるハイパーメディアデータは、他からは参照されるが、自身はリンクの終点に位置するハイパーメディアデータである。また、一般にデータは作成、更新、削除が行われることで、時間的に変化する。ここで、データのある時間的瞬間の状態夫々をバージョンと呼ぶ。そして、あるデータについて、いくつかのバージョンを記憶し、また指定された古

いバージョンを取り出し、あるいは古いバージョンに戻すなど、バージョンに関する種々の操作を行うことを「バージョン管理」と呼ぶ。

【0025】図1に、本発明の一実施形態に係るバージョン管理装置の基本的な構成を示す。本バージョン管理装置は、データ取得部101、アドレス抽出部102、アドレス記憶部103、アドレス選択部104、指定条件記憶部105、データ記憶部106、バージョン管理制御部107を備えている。

10 【0026】データ取得部101は、ハイパーメディアデータ群から、指定されたアドレスに対応するデータを取得する。アドレス抽出部102は、データ取得部101で取得されたデータからアンカーとして埋め込まれているアドレスを抽出し、そのアドレスの中から指定条件記憶部105に記憶されている指定条件を満たすものを選択し、選択されたアドレスをアドレス記憶部103に記憶させる。

20 【0027】アドレス記憶部103は、アドレスおよびアドレスで指定されたデータを取得した時刻などのアドレス付加情報を記憶する。アドレス選択部104は、アドレス記憶部103に記憶されるアドレスから次にデータを取得するアドレスを選択する。アドレスの選択は、指定条件記憶部105に関係する指定条件がある場合は、該指定条件に基づいて行われ、指定条件がない場合は、アドレス付加情報などを利用して選択する。

30 【0028】指定条件記憶部105は、あるアドレスのデータ取得時間間隔や取得するアドレスの範囲などの条件である指定条件を記憶する。データ記憶部106は、データ取得部101により取得したデータまたはそのデータを加工したデータ、および対応するアドレスを、そのデータの取得時刻などのバージョン情報を付加して記憶する。

【0029】バージョン管理制御部107は、データ記憶部106に記憶されている情報を使って、指定されたバージョンのデータを取り出すなどのバージョン管理にかかわる制御を行う。

40 【0030】本実施形態では、以下のような手順により処理を行う。まず、アドレス記憶部103に起点アドレスを記憶させる。起点アドレスは、ハイパーメディアデータをアンカーから辿って次々にデータを取得していくときの起点となるアドレスである。従来のバージョン管理装置ではすべてのアドレスを指定しなければならないが、本実施形態では起点アドレスのみを指定すれば良い。なお、起点アドレスとしては、1つのアドレスを設定しても良いし、複数のアドレスを設定しても良い。

50 【0031】次に、アドレス選択部104が、アドレス記憶部103に記憶されているアドレスから1つを選択する。選択にあたっては、指定条件記憶部105に記憶されている指定条件が参照される。例えば、まだ一度もデータを取得していないアドレスや以前に取得した時刻

が古いものなどを優先的に選択する方法もある。

【0032】次に、アドレス選択部104において選択されたアドレスで指定されるハイパーメディアデータの取得を、データ取得部101が行う。取得されたデータは、アドレス抽出部102によりアンカーとして埋め込まれているアドレスが抽出され、抽出されたアドレスの中から指定条件記憶部105に記憶される指定条件を参照して記憶されるべきアドレスを選択し、アドレス記憶部103に記憶させる。

【0033】一方、該取得されたデータはそのままの形で、または加工された後、データの取得時刻などのバージョン情報を付加してデータ記憶部106に記憶させる。このとき、過去のバージョンは無変更のままか、または他の保存されているデータから再構成可能な形に加工される。

【0034】以降、アドレス選択部104によるアドレスの選択に戻り、データの取得等が繰り返される。以上の手続により、本実施形態では、管理すべきハイパーメディアデータすべてを指定することなく、一部のアドレスを指定するだけで次々にデータを取得して各バージョンを保存し、バージョン管理を行うことが可能になる。

【0035】また、図1には示されていないが、本実施形態では、検索サービス部を持つ場合がある。検索サービス部は、データ記憶部106に記憶されるデータを対象として検索する機能を持つ。これにより、様々なバージョンのデータを検索することが可能となる。例えば、WWWで代表される大規模ハイパーメディア系の過去のバージョンまで含めた検索が可能となる。

【0036】以下、本実施形態についてより詳しく説明する。なお、以下では、本発明の一実施形態として、ハイパーメディアデータのバージョン管理を行うバージョン管理装置を例にとり、特にハイパーメディアデータの取得および検索サービスを中心に説明する。

【0037】前述のようにハイパーメディアデータとは、データ自身の中に他のデータまたは自分自身のデータを指すアドレス（アンカー）が0または1以上埋め込まれているデータである。

【0038】ハイパーメディアデータの例としては、インターネット上に展開されるWWWがあげられる。この場合、ハイパーメディアデータ群は、WWWにおいて各サーバに分散されているハイパーメディアデータの集まりであり、各データのアドレスは、例えば、ネットワーク中でサーバを特定する情報と、そのサーバ内でデータを特定する情報から構成される。

【0039】前述のように、一般に、これらデータは、作成された後、更新あるいは削除により変化して行く。そこで、過去のデータを保存しておき、後に過去のデータを取り出したり、どのように変化したかを調べたりするバージョン管理を行うことが有効となる。

【0040】本実施形態のバージョン管理装置は、概要

としては、ハイパーメディアデータに埋め込まれているアドレスを次々に辿りながら、対応するハイパーメディアデータを取得し、そのデータにバージョン情報を付加して保存したものを使ってバージョン管理を行う。このため、管理したいハイパーメディアデータをあらかじめすべて指定する必要がなく、最低限の指定をしておけば、後は自動的にデータの取得を行い、保存を行う。これは、全データを把握し切れていないときなどに特に有効である。

【0041】また、本実施形態におけるもう1つの特徴として、検索サービス装置を用いることによりこれらの保存された様々なバージョンのデータを対象として検索を行うことが可能である点があげられる。これにより、例えばハイパーメディア群においてはすでに失われてしまった過去のバージョンから検索を行うことが可能となる。

【0042】図1は、本実施形態におけるバージョン管理装置の基本的な構成を示す図である。本バージョン管理装置は、データ取得部101、アドレス抽出部102、アドレス記憶部103、アドレス選択部104、指定条件記憶部105、データ記憶部106、バージョン管理制御部107を備えている。

【0043】データ取得部101は、分散して存在するハイパーメディアデータ群から、指定されたアドレスに対応するデータを取得する。アドレス抽出部102は、データ取得部101で取得されたデータからアンカーとして埋め込まれているアドレスを抽出し、アドレス記憶部103に記憶させる。ただし、抽出されたアドレスを必ずしもすべて記憶させるわけではなく、指定条件記憶部105に記憶された指定条件をもとに、記憶されるアドレスを選択する。

【0044】アドレス記憶部103は、アドレスおよびアドレス付加情報を記憶する。アドレス付加情報とは、データを取得したかを表すフラグや、データを取得した時刻など、アドレスに関係する情報のことである。

【0045】アドレス選択部104は、アドレス記憶部103に記憶されるアドレスから次にデータを取得するアドレスを選択する。アドレスの選択は、指定条件記憶部105に指定条件がある場合は、該指定条件に基づいて行われる。指定条件がない場合は、まだデータを取得していないアドレスや、以前のデータの取得時刻が最も古いアドレスなどが優先して選択される。

【0046】指定条件記憶部105は、あるアドレスのデータ取得時間間隔や取得するアドレスのグループなどの指定条件を記憶する。データ記憶部106は、データ取得部101により取得したデータまたはそのデータを加工したデータ、および対応するアドレスを、バージョン情報を付加して記憶する。バージョン情報とは、そのデータを取得した時刻や同一アドレスのときは何度目に記録されたかを表すバージョン番号などの情報である。

【0047】なお、本実施形態では、新たにデータを追加するとき、旧データ（新データの直前のバージョンのもの）を新データと旧データとの差分に置き換え、新データをそのままの形で追加して記憶するものとする（すなわち、旧データに関しては加工していることになる）。

【0048】なお、上記の代わりに、各バージョンのデータを加工せずにそのまま記憶する方法、あるいは新データの方を差分化して記憶する方法、データを圧縮して記憶する方法など、種々の方法に変更することも可能である。また、用途によっては、必ずしも元のデータを再現できるように保存しておく必要はない。例えば、保存したデータを検索のみに使うのであれば、検索に都合の良い形に加工して記憶させておけば良い。

【0049】データ管理制御部107は、データ記憶部106に記憶される情報を使って、指定されたバージョンのデータを取す。その際、上記のように差分や圧縮の技術を用いる場合、2つのバージョンのデータを比較して差分を取り、あるいは指定されたバージョンのデータを復元するなど、バージョン管理にかかわる制御を行う。なお、データ管理制御部107における制御は、従来のバージョン管理技術（UNIXにおけるSCCSやRCSなど）と基本的には同様であるので、ここでの詳細な説明は省略する。

【0050】次に、バージョン管理装置内の各記憶部における記憶構造を詳しく説明する。まず、アドレス記憶部103における記憶構造について説明する。図2に、アドレス記憶部103における記憶構造の一例を示す。

【0051】本実施形態では、図2のように、アドレス（501）およびアドレス付加情報（取得状態（502）、取得時刻（503）、取得所要時間（504）、エラー番号（505）を1単位とし、リスト構造で管理する。

【0052】ここで、アドレス（図中501）とは、ハイパーメディアデータを指し示すアドレスそのものである。取得状態（図中502）とは、アドレスに対応するデータの取得に関する状態であり、「未取得」「取得成功」「取得失敗」の3状態のいずれかが記憶される。

【0053】「未取得」とは、まだそのアドレスに対応するデータの取得が試行されていない状態である。「取得成功」とは、最後に行なったデータ取得が成功したことを示す。このとき、取得時刻（図中503）にはデータを取得した時刻が、取得所要時間（図中504）にはデータを取得するのに要した時間が記憶される。

【0054】「取得失敗」とは、最後に行なったデータ取得が失敗したことを示す。このとき、失敗した理由がエラー番号（図中505）として記録される。失敗の理由としては、（1）タイムアウト、（2）そのアドレスに対応するデータが存在しなかった、（3）セキュリティ上の理由などでデータの取得が許可されなかった、

などがあり、それぞれが対応するエラー番号を持つ。このとき、取得時刻（図中503）には取得を試みた時刻が、取得所要時間（図中504）には取得を試みてから失敗の反応が返ってくるまでの時間が記録される。

【0055】なお、ここまでは取得状態として直前の状態のみを記憶しているが、アドレス選択部104などでさらに細かな選択の調整をするときは、もっと以前の状態も併せて記録しておけば、それを利用することも可能となる。

10 【0056】また、本実施形態では、アドレス記憶部103における記憶構造としてリスト構造を採用しているが、ハッシュ構造などに置き換えてアクセスの効率を上げることも可能である。

【0057】次に、指定条件記憶部105における記憶構造を説明する。指定条件記憶部105では、テキストとして指定条件を記憶する。テキストである理由は、指定条件は管理者など人間が直接扱う可能性が大きいためである。なお、テキスト以外の記録構造もちろん可能である。

20 【0058】図3に、指定条件記憶部105における記憶構造の一例を示す。起点アドレスとは、ハイパーメディアデータをアンカーから辿って次々にデータを取得していくときの起点となるアドレスである。起点アドレスは一般的に複数のアドレスであり（もちろん、1つのアドレスでも構わない）、処理の最初にアドレス記憶部103に記憶される。

30 【0059】ここで、「取得予定時刻」とは、あるアドレスで指し示されるデータを取得する時刻を既定する。時刻の他に曜日などの指定もでき、例えば「毎週水曜日の3時に取得せよ」などといった指定を実現するための記述が可能である。なお、取得予定時刻はアドレスの範囲について規定することが可能である。

40 【0060】アドレス範囲とは、取得対象とするアドレスの範囲を表すものであり、1つのアドレス、または複数のアドレスを、[] で区切って並べたもの、または「アドレス1ーアドレス2」のように、[-] を使って直接範囲を指定したもの、またはこれらを組み合わせで指定することができる。また、その他に、[and]、[or]、[not] などを使った論理演算、および[?]、[?] などを使ったワイルドカード指定を用いることも可能である。アドレス範囲を表す式の解釈は、従来の構文解析技術を使うことが可能である。

【0061】「取得時間間隔」とは、あるアドレスについて、対応するデータが取得されてから次に取得されるまでの時間間隔を指定する。取得時間間隔はアドレス範囲に対して指定することができる。

50 【0062】「取得アドレス範囲」とは、あるアドレスの範囲を指定したものである。すなわち、ここに記述されていない範囲のアドレスは取得されない。なお、この欄がなかったときは、本実施形態においては全アドレス

を取得対象とするものとする。

【0063】次に、データ記憶部106における記憶構造について説明する。図4は、データ記憶部106における記憶構造の一例を示す。各バージョンのデータはアドレスごとにファイルに分けられて記録される。

【0064】ファイル名は、例えば一部にアドレス名を含むなど、アドレスから一意に対応付けられるものとする。例えば、図中701のように、アドレスnのデータのファイル名を、アドレスnファイルとする。

【0065】各ファイルには、まずアドレス（図中711）が記憶され、次に各バージョンに対応する情報が並ぶ。各バージョンは、バージョン情報（取得時刻（712）、記憶時刻（713）、サイズ（714）、タイプ（715）、バージョン番号（716）、コメントサイズ（717）、コメント（718））とデータ部（差分フラグ（719）、差分データサイズ（720）、差分データ（721））を持つ。

【0066】「取得時刻」とは、対応するバージョンのデータがデータ取得部101に取得された時刻を表す。

「記憶時刻」とは、対応するバージョンにおける各情報がデータ記憶部106に記憶された時刻を表す。記憶時刻は、一般に、取得時刻に近い時刻となる。

【0067】「サイズ」とは、対応するバージョンのデータの実サイズを表す。「タイプ」とは、対応するバージョンにおけるデータのタイプを表す。タイプには、例えばテキスト、静止画、音声、動画などがある。当初テキストだったものが後で静止画に変わるなど、バージョンによってタイプは変化する可能性があるため、各バージョンごとにタイプを記憶できるようにしている。

【0068】「バージョン番号」とは、対応するバージョンに一意に付けられた番号である。特に指定しない限り、バージョンを追加するごとに1ずつ増加するものとする。

【0069】「コメントサイズ」とは、コメントのサイズを表す。「コメント」とは、対応するバージョンに関するコメントであり、任意のコメントを挿入することができる。

【0070】「差分フラグ」とは、対応するバージョンの記憶されるデータが差分化されているか否かを表すフラグである。本実施形態では、通常は、対応するアドレスの最後のバージョンを除いて差分フラグはON（差分化される）である。また、このフラグにより差分化するかどうかを指定することも可能である。

【0071】「差分データサイズ」とは、差分化されたデータのサイズである。「差分データ」とは、データ本体を記憶する部分である。本実施形態では、差分フラグがONのときは次のバージョンとの差分をとったデータが記憶され、差分フラグがOFFのときはそのままのデータが記憶される。本実施形態では、対応するアドレスの最後のバージョンのデータは、必ず差分化されてい

いデータが記憶される。

【0072】なお、データが現実更新された時刻を知り得る場合には、バージョン情報に更新された時刻を含めても良い。以下では、本実施形態のバージョン管理装置におけるデータ取得処理の流れについて説明する。

【0073】図5は、本バージョン管理装置の処理の流れを示すフローチャートである。最初にステップS1で全体の初期化を行う。例えば、記憶されている情報をクリアする。

10 【0074】次に、ステップS2でアドレス記憶部103に起点アドレスを記憶させる。前述したように起点アドレスとは、ハイパーメディアデータを順に辿ってデータを取得するとき、起点となるアドレスのことである。起点アドレス記憶時は、アドレス付加情報である取得状態を「未取得」にする。

【0075】次のステップS3からはループ処理に入る。ステップS3では、アドレス選択部104が次に取得すべきデータのアドレスをアドレス記憶部103から選択する。

20 【0076】この選択にあたって、指定条件記憶部15に記憶されている情報を参照する。例えば、アドレス記憶部103にアドレスAが記憶されていたとしても、指定条件記憶部105に記憶される取得アドレス範囲にアドレスAが含まれていないときは、アドレスAは選択されない。また、指定条件記憶部105に、アドレスBの取得時間間隔が1時間であると指定されていて、アドレス記憶部103に記憶されるアドレスBの以前の取得時刻が現在から1時間より前であったとき、アドレスBが優先して選択する。

30 【0077】また、例えば、あるアドレスについて、取得予定時刻と取得時間間隔の両方で指定がなされているような場合は、いずれか一方でも条件を満たせば選択する方法、両方の条件を満たした場合に選択する方法、予め定めた方を優先適用し、他方を無視する方法などが考えられる。

【0078】なお、指定条件記憶部105に特に指定がないときは、まだデータを取得していないアドレスや、以前に行われたデータの取得時間が最も古いアドレスなどが優先して選択される。

40 【0079】また、取得状態が「取得失敗」であり、取得を試みた時刻が最近だった場合などは、そのアドレスを選択されにくいようにしても良い。ステップS4では、ステップS3におけるアドレスの選択が成功したか否かを判断する。アドレスの選択に失敗した場合は、指定条件記憶部105の指定条件に対して、アドレス記憶部103に記憶されているアドレスがすべて選択不可能の状態である。例えば、すべてのアドレスがまだ取得予定時刻に達していない状態などがそれに該当する。そのような場合は、選択可能状態になるまで待機する（ステップS5）。

【0080】ステップS3でのアドレスの選択に成功した場合、ステップS6では、選択されたアドレスに対応するデータをデータ取得部101により取得する。ステップS7では、データの取得が成功したか否かを判断する。

【0081】データの取得に失敗したときは、後でアドレス選択の一条件として利用するために、アドレス記憶部103でそのアドレスの取得状態を「取得失敗」とし、失敗理由をエラー番号として書き込む（ステップS8）。そして、ステップS3に戻る。

【0082】一方、ステップS3でのデータの取得に成功したときは、ステップS9に進む。ステップS9では、アドレス記憶部103において、対応するアドレスのアドレス付加情報を更新する。具体的には、取得状態を「取得成功」とし、その他にデータ取得時刻、データ取得所要時間などを記憶させる。

【0083】次のステップS10では、ステップS6で取得したデータを、バージョン情報を付加してデータ記憶部16に記憶する。なお、前述したように本実施形態では、前バージョンのデータを現バージョンのデータとの差分で置き換えて、現バージョンのデータをそのまま記憶させている。ここで、差分の計算自体は、従来の技術（UNIXにおけるdiffコマンドなど）を用いることで可能である。なお、前のバージョンのデータとの差分がなかったとき、すなわち変化がなかったときは、取得時間等のバージョン情報は記憶させるが、バージョン番号は変化させないものとする。

【0084】次のステップS11では、ステップS6で取得したデータにアンカーとして埋め込まれているアドレスを抽出し、指定条件記憶部105に記憶される指定条件を参照して、抽出されたアドレスのうちから該指定条件を満足するものを選択し、選択されたアドレスをアドレス記憶部103に記憶させる。例えば、指定条件記憶部105に取得アドレス範囲が記憶されていたときは、抽出されたアドレスの中からその範囲に含まれるアドレスのみを選択し、アドレス記憶部103に記憶させる。指定条件記憶部105に特に関係する指定条件がないときは、そのまますべてを記憶させる。なお、記憶させるにあたって対応するアドレスの取得状態を「未取得」とする。また、アドレス記憶部103に同一のアドレスがすでに記憶されているときは、該アドレスについては新たに変更等は行わない。

【0085】この後、ステップS3に戻り、ループ処理を繰り返す。以上が各ハイパーメディアデータを取得し、バージョンを付加して記憶させるための手続である。

【0086】以上のように本実施形態によれば、取得されたハイパーメディアデータにアンカーとして埋め込まれているアドレスにより指し示されるハイパーメディアデータを抽出して連鎖的に取得対象としていくので、バ

ージョン管理すべきハイパーメディアデータのアドレスすべてを指定することなく一部のアドレスを指定するだけで次々にハイパーメディアデータを取得することができるとともに、所定のタイミングで繰り返し各ハイパーメディアデータを取得することができ、各タイミングで取得された各ハイパーメディアデータについてのバージョン管理が可能となる。

【0087】これにより、各ハイパーメディアデータについて様々なバージョンの検索が可能となる。次に、図10のバージョン管理装置に付加する検索サービス装置の一実施形態について説明する。

【0088】本実施形態では、テキストデータを検索の対象として説明するが、本発明はテキストデータに限定されず、例えば、静止画や音声など他のメディアデータについてパターンマッチング法など従来の技術を利用して検索対象にすることも可能である。

【0089】本実施形態では、検索の前段階として、各バージョンのテキストデータからそれに含まれる単語を抽出し、データベース化を行う。そして、検索サービスでは、検索キーとなる単語やバージョン番号の範囲や更新日の範囲などの指定を受け、これらをもとに検索データベースとの間でマッチングを行い、その結果を検索結果とする。

【0090】図6は、本実施形態に係る検索サービス装置の構成を示す図である。本検索サービス装置は、検索データベース作成部202、検索データベース記憶部203、検索制御部204を備えている。

【0091】図6中のデータ記憶部は、図1に示したデータ記憶部106である。検索データベース作成部202は、データ記憶部106に記憶されるデータから検索用のデータベースを作成する。

【0092】検索データベース記憶部203は、検索用データベースを記憶する。検索制御部204は、図示しない検索式入力部から入力された検索式に基づき、検索データベース記憶部203に記憶されるデータベースを参照して検索を行う。

【0093】次に、検索データベース記憶部203における記憶構造について説明する。図7に、検索データベース記憶部203における記憶構造の一例を示す。本実施形態の検索データベースは、単語をキーとしてその単語を含むデータのアドレスとバージョンに関する配列を引く、ハッシュ構造になっている。

【0094】図7において、単語（801）をキーにして、アドレス（811）とバージョン番号（812）と付加情報（813）を要素とする配列を引く。アドレスとは、その単語を含むデータのアドレスである。

【0095】「バージョン番号」とは、対応するアドレスの中のバージョンを表す番号であり、図4中のバージョン番号716に対応している。「付加情報」とは、最低限必要な検索結果であるアドレスとバージョン番号に

加える、付加的な情報を提供するものであり、例えば、更新時刻やそのバージョンのコメントなどである。なお、付加情報に更新時刻を含める場合、現実の更新時刻がわからなければ、前述の取得時刻あるいは記憶時刻で代用しても良い。

【0096】次に、本実施形態の検索サービス装置における処理の流れについて説明する。本検索サービス装置においては、大きく2つの処理が存在する。1つは、検索データベースを作成する処理で、もう1つは、実際に検索を行う処理である。

【0097】まず、検索データベースの作成について説明する。検索データベース作成部202は、データ記憶部106に記憶される各バージョンのデータからメディアのタイプを参照することでテキストデータを選択し、そのデータに対して形態素解析を行う。ここで、形態素解析とは文章をそれを構成する単語に分解することであり、機械翻訳に利用されるなどすでに従来技術として良く知られている。次に、検索データベース作成部202は、形態素解析結果に含まれる全単語を抽出し、各単語が含まれるデータに対応するアドレス、バージョン番号、および付加情報を、データ記憶部106から参照して検索データベース記憶部203に記憶させる。これにより検索データベースが作成される。

【0098】次に、外部からの検索要求対応の処理を説明する。外部からの検索要求は、検索制御部204によって処理される。検索要求は、検索キーとなる単語およびバージョン番号の範囲、更新日時の範囲などを検索制御部204に与えることでなされる。検索制御部204は、検索要求を受けると、まず検索要求に含まれる単語に関して、その単語を含むデータのアドレス、バージョン情報、および付加情報を検索データベース記憶部203から取得する。次に、取得した情報から検索要求に含まれるバージョン番号の範囲または更新日時の範囲に該当するものを選択し、それを検索結果として要求元に送る。

【0099】以下では、図1の構成と図6の構成を備えるバージョン管理装置の一例と、その利用例について説明する。図8に、図1の構成と図6の構成を備えるバージョン管理装置2の構成を示す。

【0100】本バージョン管理装置2は、バージョン記憶部10、バージョン管理制御部20、検索部30を備えている。バージョン記憶部10は、図1のデータ取得部101、アドレス抽出部102、アドレス記憶部103、アドレス選択部104、指定条件記憶部105、データ記憶部106を有する。

【0101】バージョン記憶部10は、既に説明したような手順で、ハイパーメディアデータ群（図中4）からデータを取得し、バージョン情報を付加してデータ記憶部106に記憶する。

【0102】ここで、ハイパーメディアデータ群とは、

各データがアドレス付けされたハイパーメディアデータの集合である。例としてはインターネット上のWWW (World Wide Web) がある。各ハイパーメディアデータは、一般にネットワークを介して分散して配置されている。

【0103】バージョン管理制御部20は、図1のバージョン管理制御部107を有する。バージョン管理制御部20は、既に説明したような手順で、ユーザの要求により過去のバージョンを取り出したり、異なるバージョン間での比較を行うなど、データのバージョンを管理する。

【0104】検索部30は、図6の検索データベース作成部202、検索データベース記憶部203、検索制御部204を有する。検索部30は、既に説明したような手順で、データ記憶部106に記憶されたバージョン情報からデータベースを構成し、ユーザの要求により検索を行う。

【0105】ユーザは、通常、ユーザ端末8から取得したいハイパーメディアデータのアドレスをデータ取得部6に指定することで、ハイパーメディアデータ群から該当データを取得することができる（図8中a）。これは、従来の取得方法であり、ハイパーメディアデータは時間的に変化していることから、最新のバージョンのみを取得することができる。

【0106】本実施形態では、ハイパーメディアデータの様々なバージョンを記憶していることから、過去のバージョンを取り出したり、検索したりすることが可能となる。

【0107】本実施形態では、検索部30自身がハイパーメディアデータと同様なアドレスを持つものとする（検索部30を1つのサーバ装置として扱う）。ユーザは、ユーザ端末8を用いて、検索キーや検索対象となるバージョンなどを検索部30のアドレスに付加し、データ取得部6に指定する。すると、検索部30が検索を行い、ユーザはその結果を取得することができる（図8中b）。

【0108】また、本実施形態では、バージョン管理制御部20自身もハイパーメディアデータと同様なアドレスを持つものとする（バージョン管理制御部20を1つのサーバ装置として扱う）。ユーザは、ユーザ端末8を用いて、取得したいバージョンをバージョン管理制御部20のアドレスに付加してデータ取得部6に指定することで、そのバージョンのデータを取得することができる（図8中c）。これにより、現在のハイパーメディアデータを閲覧するのと同様に過去のバージョンのデータを閲覧することが可能となる。

【0109】バージョンの指定方法としては、明示的にバージョン番号を指定する方法の他に、日時またはその範囲を指定する方法がある。後者の場合、バージョン管理装置では、指定された日時またはその範囲に該当する

データを提供する。なお、日時で指定された場合、指定された日時と同じ更新時刻（あるいは取得時刻または記憶時刻）の情報を持つものだけではなく、（１）指定された日時以前で最も新しい取得時刻または記憶時刻を持つものも提供する、（２）指定された日時以降で最も古い取得時刻または記憶時刻を持つものも提供する、あるいは（３）上記の（１）と（２）を両方提供するようにしても良い。本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して実施することができる。

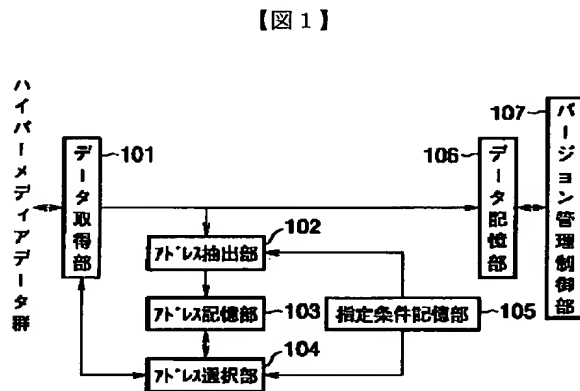
【0110】

【発明の効果】本発明によれば、取得されたハイパーメディアデータにアンカーとして埋め込まれているアドレスにより指し示されるハイパーメディアデータを抽出して連鎖的に取得対象としていくので、バージョン管理すべきハイパーメディアデータのアドレスすべてを指定することなく一部のアドレスを指定するだけで次々にハイパーメディアデータを取得することができるとともに、所定のタイミングで繰り返し各ハイパーメディアデータを取得することができ、各タイミングで取得された各ハイパーメディアデータについてのバージョン管理が可能となる。これにより、各ハイパーメディアデータについて様々なバージョンの検索が可能となる。

【図面の簡単な説明】

【図１】本発明の実施の形態に係るバージョン管理装置の基本構成を示す図

【図２】アドレス記憶部における記憶構造の一例を示す図



* 【図３】指定条件記憶部における記憶構造の一例を示す図

【図４】データ記憶部における記憶構造の一例を示す図

【図５】同実施の形態における処理の流れを示すフローチャートである。処理の流れ図

【図６】検索サービス装置の構成例を示す図

【図７】検索データベース記憶部における記憶構造の一例を示す図

【図８】同実施の形態に係るバージョン管理装置の構成

例を示す図

【符号の説明】

2…バージョン管理装置

4…ハイパーメディアデータ群

8…ユーザ端末

10…バージョン記憶部

20…バージョン管理制御部

30…検索部

101…データ取得部

102…アドレス抽出部

20 103…アドレス記憶部

104…アドレス選択部

105…指定条件記憶部

106…データ記憶部

107…バージョン管理制御部

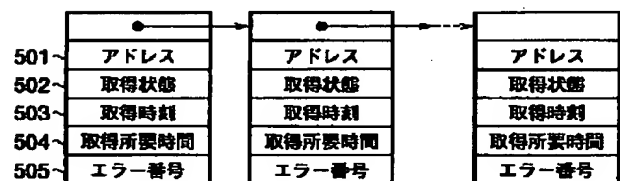
202…検索データベース作成部

203…検索データベース記憶部

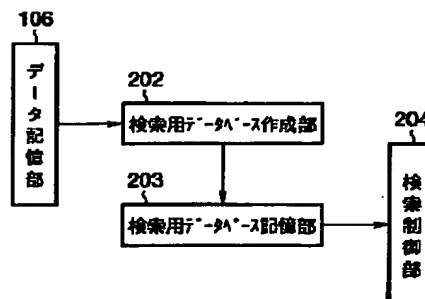
204…検索制御部

*

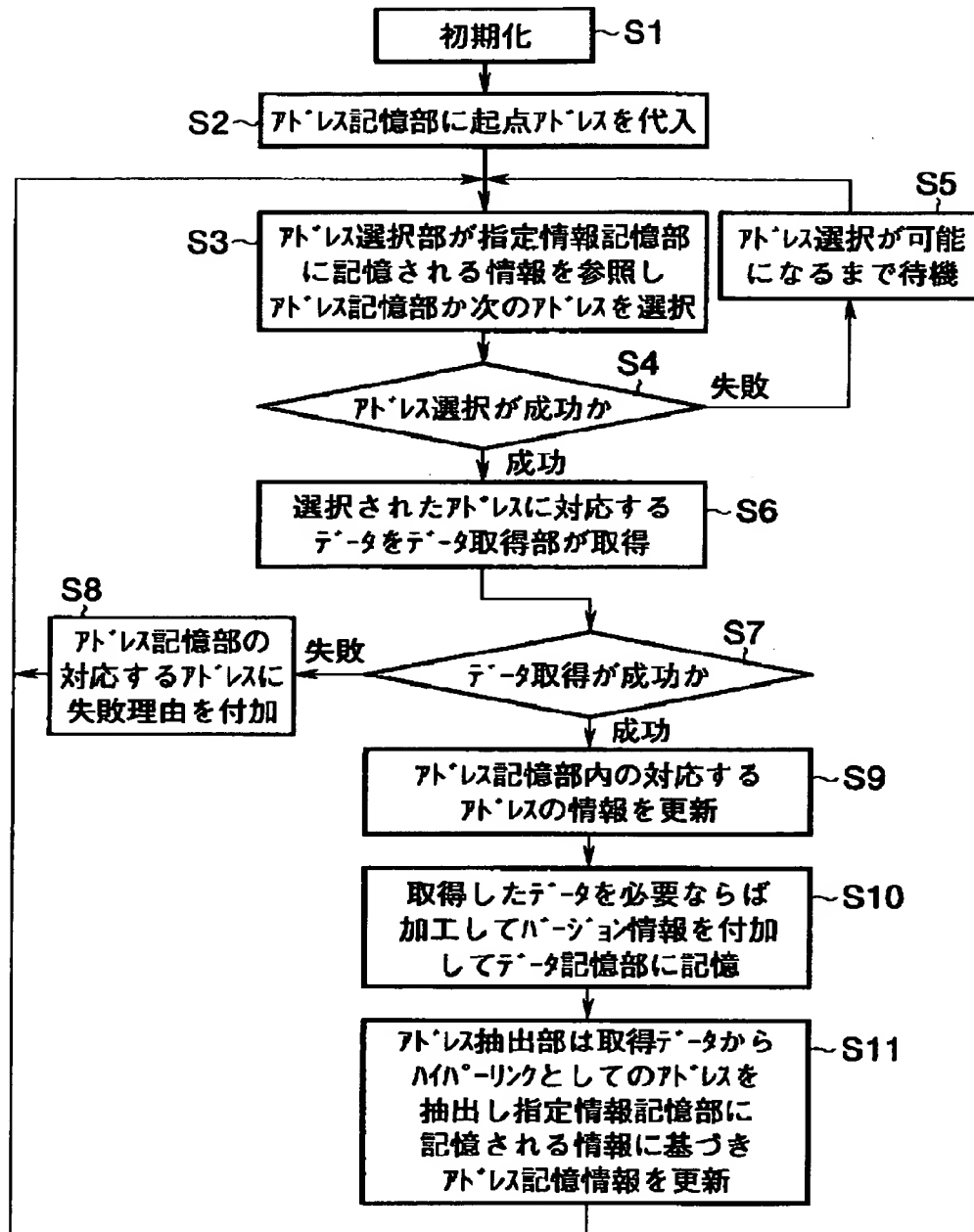
【図２】



【図６】



【図5】



【図8】

